4-1-2024

Metabolite signatures associated with microRNA miR-143-3p serve as drivers of poor lung function trajectories in childhood asthma

Kevin M. Mendez
Edith Cowan University

Sofina Begum

Anshul Tiwari

Rinku Sharma

Qingwen Chen

See next page for additional authors

10.1016/j.ebiom.2024.105025
Mendez, K. M., Begum, S., Tiwari, A., Sharma, R., Chen, Q., Kelly, R. S., . . . Lasky-Su, J. A. (2024). Metabolite signatures associated with microRNA miR-143-3p serve as drivers of poor lung function trajectories in childhood asthma. eBioMedicine, 102, article 105025. https://doi.org/10.1016/j.ebiom.2024.105025

Metabolite signatures associated with microRNA miR-143-3p serve as drivers of poor lung function trajectories in childhood asthma

Kevin M. Mendez,[a,b] Sofina Begum,[a] Anshul Tiwari,[a,c] Rinku Sharma,[a] Qingwen Chen,[a] Rachel S. Kelly,[a] Nicole Prince,[a] Mengna Huang,[a] Priyadarshini Kachroo,[a] Su H. Chu,[a] Yulu Chen,[a] Kathleen Lee-Sarwar,[a,d] David I. Broadhurst,[b] Stacey N. Reinke,[b] Robert Gerszten,[e] Clary Clish,[f] Lydiana Avila,[g] Juan C. Celedón,[g] Craig E. Wheelock,[h,i] Scott T. Weiss,[a] Michael McGeachie,[a,j] and Jessica A. Lasky-Su[a,*,j]

[a]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[b]Centre for Integrative Metabolomics & Computational Biology, School of Science, Edith Cowan University, Perth, Australia
[c]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA
[d]Division of Allergy and Clinical Immunology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[e]Department of Cardiovascular Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA
[f]Broad Institute, Cambridge, MA, USA
[g]Division of Pediatric Pulmonary Medicine, UPMC Children's Hospital of Pittsburgh, University of Pittsburgh, Pittsburgh, PA, USA
[h]Unit of Integrative Metabolomics, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[i]Department of Respiratory Medicine and Allergy, Karolinska University Hospital, Stockholm, Sweden

eBioMedicine
2024;102: 105025
Published Online 7 March 2024
https://doi.org/10.1016/j.ebiom.2024.105025

Summary

Background Lung function trajectories (LFTs) have been shown to be an important measure of long-term health in asthma. While there is a growing body of metabolomic studies on asthma status and other phenotypes, there are no prospective studies of the relationship between metabolomics and LFTs or their genomic determinants.

Methods We utilized ordinal logistic regression to identify plasma metabolite principal components associated with four previously-published LFTs in children from the Childhood Asthma Management Program (CAMP) (n = 660). The top significant metabolite principal component (PC_{LF}) was evaluated in an independent cross-sectional child cohort, the Genetic Epidemiology of Asthma in Costa Rica Study (GACRS) (n = 1151) and evaluated for association with spirometric measures. Using meta-analysis of CAMP and GACRS, we identified associations between PC_{LF} and microRNA, and SNPs in their target genes. Statistical significance was determined using an false discovery rate-adjusted Q-value.

Findings The top metabolite principal component, PC_{LF}, was significantly associated with better LFTs after multiple-testing correction (Q-value = 0.03). PC_{LF} is composed of the urea cycle, caffeine, corticosteroid, carnitine, and potential microbial (secondary bile acid, tryptophan, linoleate, histidine metabolism) metabolites. Higher levels of PC_{LF} were also associated with increases in lung function measures and decreased circulating neutrophil percentage in both CAMP and GACRS. PC_{LF} was also significantly associated with microRNA miR-143-3p, and SNPs in three miR-143-3p target genes; CCZ1 (P-value = 2.6×10^{-5}), SLC8A1 (P-value = 3.9×10^{-5}); and TENM4 (P-value = 4.9×10^{-5}).

Interpretation This study reveals associations between metabolites, miR-143-3p and LFTs in children with asthma, offering insights into asthma physiology and possible interventions to enhance lung function and long-term health.

Funding Molecular data for CAMP and GACRS via the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung, and Blood Institute (NHLBI).

Keywords: Asthma; Lung function trajectories; Metabolomics; MicroRNAs; Genomics

*Corresponding author. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, 181 Longwood Avenue, Boston, MA 02115, USA.
E-mail address: rejas@channing.harvard.edu (J.A. Lasky-Su).
[j]Authors contributed equally to this work.

Research in context

Evidence before this study

Lung function trajectories (LFTs) in asthma critically influence long-term outcomes, marking a continuum from stable development to early decline, with direct implications for chronic obstructive pulmonary disease (COPD) risk. Deciphering these trajectories is key in addressing childhood asthma's long-term impact, with evidence indicating genetics play a substantial role in LFT outcomes.

Added value of this study

Here we identified a metabolomic signature linked with LFTs in children with asthma, a previously unexplored area. By uncovering the association between specific plasma metabolites and LFTs, as well as their genomic determinants, the research provides an insight into the pathophysiology of asthma. These associations in a pediatric cohort offers a valuable perspective for early intervention strategies aimed at improving long-term respiratory health outcomes in individuals with asthma.

Implications of all the available evidence

Our study revealed that a range of metabolites, including exogenous and microbial-derived compounds, together with miR-143-3p, are associated with LFTs in children with asthma, suggesting dietary interventions may influence these diverse metabolic pathways for therapeutic gains.

Introduction

Asthma imparts a significant global public health burden, affecting approximately 500 million people worldwide, with the majority of cases originating in early life and contributing to substantial healthcare costs, morbidity, and reduced quality of life.[1,2] Spirometric lung function is a critical facet of asthma and other chronic lung diseases and is an important indicator of asthma severity, progression, and prognosis.[3] While current clinical use of lung function has primarily focused on spirometry at the time of visit, lung function trajectory (LFT) is more important for long-term health, as this measure identifies individuals at the greatest risk of poor long-term respiratory health. McGeachie et al.[4] classified LFT growth patterns from childhood to early adulthood into distinct groups based on the severity of lung function decline and reduced lung growth throughout development. These groups, listed here in increasing order of severity, are normal growth (NG), early decline (ED), reduced growth (RG), and reduced growth with early decline (RG/ED).[4] Critically, individuals in the RG and RG/ED groups were associated with a heightened risk of chronic obstructive pulmonary disease (COPD) compared to NG (P-value < 0.001) (determined via exact Mantel-Haenszel chi-square test for trend), with 16% of RG and 21% of RG/ED showed signs of COPD, in stark contrast to only 1% in NG and 5% in the ED group.[4] LFT is a product of both genetic and environmental influences. Further understanding of the major drivers of LFT may provide insight into understanding and prevention of poor long-term respiratory health among individuals with childhood asthma.

Metabolomics, described as the "chemical fingerprint" of cellular processes, investigates the metabolites within an organism, offering insights into its current health state.[5] These metabolites, ranging from amino acids to lipids, serve as indicators of biochemical activities and the interplay between genetics and environment. In asthma research, metabolomics has identified unique metabolic profiles associated with different phenotypes, highlighting pathways not previously linked to asthma pathogenesis.[6,7] Discoveries in altered lipid metabolism and oxidative stress, for example, have enriched our understanding of asthma's molecular foundation and have hinted at new diagnostic and treatment methods.[8,9] With the progression in metabolomics, personalized medicine for asthma, customized based on an individual's metabolic makeup, is becoming more feasible.[8,9]

There is a growing body of metabolomic studies on asthma diagnosis and lung function.[6] Yet, there remains a gap in dedicated, prospective studies examining the association between specific metabolites and LFTs, a crucial aspect for understanding long-term lung health. Lung function development hinges on the intricate balance between genetic factors and environmental exposures. Early life, a phase characterized by rapid lung growth and augmented plasticity, is especially vulnerable to environmental risk factors, which can culminate in chronic respiratory morbidities persisting into adulthood.[10] Factors spanning from in utero environments, nutrition, and tobacco exposure to early life infections, microbiome variations, and allergies can significantly mold lung functional development.[10] Some determinants, like prematurity, exert a direct impact on lung growth, while many, such as recurrent infections, channel their influence via inflammatory or oxidative stress pathways. Intricately, these environmental exposures leave discernible metabolic traces, detectable through metabolomics. Thus, delving into metabolomic profiles can not only shed light on the genetic factors influencing LFTs but also pinpoint crucial environmental factors. Such insights could pave the way for targeted interventions to enhance long-term respiratory health. Augmenting this, integrating data from other

'omics spheres, especially genetics and microRNA (miR), can refine our understanding of the intricate interplay between genetics and environmental factors in dictating lung health.

MicroRNAs (miRs), as gene expression regulators, play significant roles in the genome transcription process and influence critical asthma-related pathways, such as inflammatory cascades and chronic airway obstruction.[11–13] Complementing this, genetic studies have identified numerous disease-linked variations, underscoring the complex etiology of asthma.[14–16] Given that the metabolome provides a real-time snapshot of the interplay between genome and environment, the integration of miRs and genetics with metabolomics offers a comprehensive, holistic approach. This strategy can help to unravel the drivers of different LFTs and may open new avenues for therapeutic strategies to mitigate the extensive health implications of asthma.

The primary objective of this study was to identify metabolites associated with LFTs using a well-characterized childhood asthma cohort with over 15 years of long-term follow-up. We identified a distinct metabolomic signature associated with LFTs and identified associated miRs and their gene targets to further understand the potential etiology of these LFT-driven metabolites.

Methods

Study populations

The Childhood Asthma Management Program (CAMP) (Clinicaltrials.gov; register: NCT00000575) was designed as a multi-center, randomized, double-masked, clinical trial.[17] CAMP operated as a randomized, placebo-controlled trial, focusing on the administration of budesonide, nedocromil, or placebo for children with mild-to-moderate asthma. It subsequently transitioned into three phases of observational follow-up. For inclusion, asthma was defined by the presence of 2 or more symptoms per week, utilizing an inhaled bronchodilator at least twice weekly or daily asthma medication, and showcasing airway responsiveness to methacholine <12.5 mg/ml. This trial recruited 1041 children aged 5–12 years with mild-to-moderate asthma. Standard protocols were used for spirometry, white blood cell profiling, and plasma collection. Long-term spirometric follow-up over ~17 years from the ages of 5–12 to 23–30 were previously used to develop the LFT profiles in CAMP,[4] the primary outcome of interest in this study. Smoothed longitudinal Forced Expiratory Volume in 1 s (FEV_1) trajectories were previously compared to normal charts for a person of the same age, height, sex, and race and then classified into the four ordinal categories of LFT: 1) normal lung function (NG); 2) early decline (ED); defined as having decline of FEV_1 from peak plateau before age 23; 3) reduced growth (RG); defined as being below the 25th percentile of FEV_1; and 4) both ED and RG (RG/ED). FEV_1 measurements were systematically collected across multiple stages, including the screening & baseline phase, treatment phase, transition phase, and the follow-up phase (Supplementary Fig. S1). On average, there were 21.1 (with a standard deviation of 2.8) lung function measures per patient.

Additional clinical variables were collected throughout the CAMP study, including lung function measures: Forced Vital Capacity (FVC); FEV_1;FEV_1/FVC, Forced Expiratory Flow 25–75 (FEF_{25-75}); and FEF_{25-75}/FVC); as well as airway hyperresponsiveness (AHR), measured by percent concentration of methacholine required to effect a 20% reduction in FEV_1 (PC_{20}), white blood cell counts, and Imunoglobulin E (IgE).

The Genetic Epidemiology of Asthma in Costa Rica Study (GACRS) was designed as a cross-sectional cohort.[18] It was initiated as a family-based genetics study focusing on childhood asthma, comprising Costa Rican schoolchildren aged 6–14 years and their parents. The defining criteria for asthma in these children was a prior doctor's diagnosis of asthma coupled with the occurrence of at least two respiratory symptoms or asthma attacks in the year preceding their enrollment in the study. This study recruited 1151 children aged 6–14 years with mild-to-moderate asthma between February 2001 and August 2008. Standard protocols were used for spirometry and plasma collection that was completed at enrollment. Standard clinical variables collected include lung function (FEV_1, FVC, FEV_1/FVC, FEF_{25-75}, FEF_{25-75}/FVC), ICS usage, provocative dose of methacholine causing a 20% fall in FEV1 (PD_{20}), white blood cell counts, and IgE. Longitudinal spirometry, and thus LFT, was not available in GACRS.

Metabolomic profiling

Plasma metabolomic profiling of the CAMP samples, collected four years after baseline, and the GACRS cohorts, collected during study enrollment, was conducted using four complementary liquid chromatography-mass spectrometry (LC-MS) methods in a single metabolomics run. This analysis including peak identification and annotation was performed as part of the Trans Omic Precision Medicine (TOPMed) initiative, following the procedures previously described.[19] Three nontargeted LC-MS methods were conducted at the Broad Institute including: i) reversed-phase C8 chromatography/positive ion mode MS detection; ii) reversed-phase C18 chromatography/negative ion mode MS detection; and iii) hydrophilic interaction liquid chromatography/positive ion mode MS detection. A targeted negative ion mode analysis of central metabolites (sugars, sugar phosphates, organic acids, purine, and pyrimidines) using LC-MS/MS was also conducted. Metabolites with coefficient of variation (CV%) > 25% or missing > 75% were then excluded. Remaining missing values were imputed using the k-nearest-neighbor

imputation method (R package "VIM"). The metabolite data, both targeted and untargeted, were subsequently log10 transformed and pareto scaled prior to analysis. Further details are available in the supplement, as well as the list of identified metabolites (Supplementary Table S1).

microRNA (miR) profiling

Small RNA sequencing (RNA-seq) was performed on serum samples from 1134 GACRS children. The sequencing of 492 baseline CAMP samples was described previously.[20] The same protocols were used to sequence both cohorts.[20] Serum samples were stored in freezers at –80 °C at the Channing Division of Network Medicine. Small RNA-seq libraries were prepared with the Norgen Biotek Small RNA Library Prep Kit (Norgen Biotek, Therold, Canada) and sequenced on the Illumina NextSeq 500 platform. The ExceRpt pipeline was utilized for the RNA-seq data quality control (QC).[21] MiRs with fewer than five mapped reads in at least 50% of the subjects were removed from the analysis. All samples successfully passed the QC in terms of both the number of mapped reads and total reads, signifying the availability of satisfactory miR concentration. DESeq2 was then used to normalize reads by relative $\log_2$ expression.[22]

Genetic profiling

Genotyping for CAMP and GACRS was previously described.[23] Briefly, genotyping for CAMP was performed on baseline samples using either the Illumina Quad 610 and Illumina 550 microarray chips (Illumina, Inc., San Diego, CA), and genotyping for GACRS was performed the Illumina BeadStation 500G platform (Illumina Inc., San Diego CA). Genotype imputation was performed using the Michigan Imputation Server38 with the Haplotype Reference Consortium (HRC) r1.1 201,639 reference panel. After imputation, only variants with imputation quality r2 $\geq$ 0.3 and minor allele frequency (MAF) $\geq$ 0.05 were retained.

Statistical analyses

A descriptive summary of the cohort demographics for both CAMP and GACRS children was generated. The overall study objective was to identify metabolites associated with LFT and furthermore, to identify associated miRs and associated variants within their target genes. To address the high degree of collinearity commonly seen in metabolomics data, our analytic approach incorporated data reduction using principal component analysis (PCA). By orthogonalizing the multivariate data, PCA ensures the independence of each ordinary least squares regression (OLR) model.[24] Furthermore, our analysis employed sequential association testing. First, PCA utilizing levels of the 589 detected metabolites, combining both untargeted and targeted metabolomic data, was performed in CAMP, and the metabolite principal components (PCs) explaining up to 95% of the variance were extracted. Secondly, OLR models were then used to assess the association between these metabolite PCs and LFT, while adjusting for age, sex, body mass index (BMI), race, and height. The four LFTs, namely normal growth, reduced growth, early decline, and reduced growth with early decline, were labeled as 1, 2, 3, and 4, respectively. The increasing numerical value indicates a deteriorating trajectory in the lung function, serving as a clear representation of lung health. This ensured that in our ordinal logistic regression models, the association precisely captures this gradation in the severity of LFTs. Significant metabolite principal component-LFT associations were identified using a P-value threshold that adheres to the Benjamini and Hochberg (BH) false discovery rate (FDR) criterion of less than 0.05.[25] To discern metabolite loadings that were significantly associated with the identified metabolite principal component-LFT, we employed parallel analysis (R package paran version 1.5.2).[26,27] This analytical method enabled us to set a threshold based on a P-value of 0.05 to determine significant metabolite loadings. Metabolite set enrichment analyses (MetaboAnalyst v.5.0[28]) was then applied to these significant metabolite loadings in each significant principal component to further explore the metabolite pathways underlying LFT. The hypergeometric test was specified for the over-representation analysis and relative-betweenness centrality for the pathway topology analysis.

In addition to the primary analysis, we utilized logistic regression models for each LFT using a one-vs-rest approach. Specifically, we analyzed one group vs. the rest, for each LFT: NG, ED, RG, and RG + ED. In these analyses, this classification was the binary outcome, and its relationship with metabolite PCs was controlled for variables like age, sex, BMI, race, and height.

We evaluated these findings in a secondary cohort, GACRS, by recapitulating significant metabolite loadings by utilizing the 589 metabolites levels in GACRS with principal component loadings generated in CAMP. Generalized linear models (GLM) were then used to assess the association between significant metabolite loadings and asthma phenotypes in both cohorts, adjusting for age, sex, BMI, race (race adjustment for CAMP only), and height (height adjustment for lung function only). These asthma phenotype outcomes are as follows: airway hyperresponsiveness (AHR) (measured as PC_{20} in CAMP and PD_{20} in GACRS), FEV_1 percent predicted, FVC percent predicted, FEV_1/FVC, FEF_{25-75}, FEF_{25-75}/FVC, log10(IgE), log10 (Eosinophil Count), and blood neutrophil percent. We then identified associations between PC_{LF} and miRs via GLMs with significant metabolite loadings and miR as the independent and dependent variables respectively, adjusting for the same covariates. Association P-values

from CAMP and GACRS were then combined using a fixed-effect meta-analysis approach; significant associations were reported using a combined P-value that met BH FDR < 0.05, and consistent direction of effects in both cohorts.

Significant miR gene targets were extracted from validated microRNA-target databases (miRecords, miRTarBase, and TarBase) using multiMiR (release 3.15).[29] We then performed an eQTL analysis between significant metabolite loadings and the identified miR gene target SNPs in GACRS and CAMP, adjusting for age, sex, and the first 4 genome-wide PCs. Association P-values from CAMP and GACRS were then combined using a fixed-effect meta-analysis; significant associations were reported using an effective number of independent tests accounting for 80 percent of the variance (ENT80).[30,31]

Ethics

The project was reviewed and approved. CAMP was approved by the Mass General Brigham Research Committee at Brigham and Women's Hospital (Protocol#: 2000-P-001130/55) by all participating clinical centers and the data coordinating center. GACRS was approved by the institutional review board of Mass General Brigham Healthcare (Protocol#: 1999-P- 001549/29), and the Hospital Nacional De Niños of Costa Rica. Child assent and parental written consent was obtained.

Role of funders

The funding agencies did not have any role in the design and conduct of the study; collection, management, and interpretation of the data; or preparation, review, or approval of the manuscript.

Results

Study populations

As detailed in previous work, 660 CAMP subjects were classified into one of four distinct LFTs, which are listed here in increasing order of severity: normal growth (n = 163), early decline (n = 167), reduced growth (n = 156), and reduced growth with early decline (n = 174).[4] Descriptive characteristics of CAMP (n = 660) and GACRS (n = 1151) at the time of metabolomic profiling (four years past baseline in CAMP) are shown in Table 1. Both cohorts had similar sex distributions, reflective of the higher incidence of childhood asthma in boys. Notably, CAMP participants at the time of metabolomic profiling were, on average, three years older than their GACRS counterparts (mean age 13.51 vs. 9.22 years), which explains the observed differences in height and BMI between the two groups. While CAMP's racial composition was diverse with 67.4% White, 14.1% Black, 10.3% Hispanic, and 8.2% other ethnicities, GACRS was exclusively Hispanic. Clinically, CAMP exhibited slightly diminished lung function metrics, with a mean FEV1% predicted at 94.40 compared to GACRS's 98.89 and a mean FEV1/FVC of 77.92 against GACRS's 84.23. Additionally, while CAMP had higher IgE levels, eosinophil counts were marginally elevated in GACRS. Additionally, CAMP had higher $\log_{10}$ IgE levels with a mean level of 2.65, which was higher than GACRS's mean of 2.50. In contrast, for $\log_{10}$ eosinophil count, GACRS presented a slightly elevated mean value of 2.60 compared to CAMP's 2.40.

Analysis of metabolite PCs with LFT

A total of 57 PCs explained 95 percent of the variance in the metabolome and were evaluated for associations with LFT using ordinal logistic regression. In our model, LFTs were categorized as: 1 for normal growth, 2 for reduced growth, 3 for early decline, and 4 for reduced growth with early decline. Statically significant association was observed with PC 29, hereafter referred to as PC_{LF}, and LFT ($\beta = -0.13$; SE = 0.04; P-value = 6.0×10^{-4}; Q-value = 0.03) (via OLR adjusted for age, sex, BMI, race, and height) (Table 2). The coefficient $\beta = -0.13$ indicates that for each unit increase in PC_{LF}, there's an associated shift towards better LFT categories (e.g., from early decline to normal growth). In Supplementary Table S2, metabolites significantly associated with PC_{LF} based on P-value < 0.05 (via OLR adjusted for age, sex, BMI, race, and height) are detailed. For Fig. 1, to offer a more comprehensive view, we expanded our inclusion criteria to an alpha significance level of 0.20, allowing for a broader representation to elucidate pathway-specific effects more effectively. Given that PC_{LF} is associated with better LFT as it increases, positive loadings represent metabolites conducive to improved lung function, while negative loadings signify metabolites linked to deteriorated lung function. Overall, PC_{LF} exhibited higher loadings from both endogenous and exogenous metabolites, with a considerable proportion of microbial-derived metabolites (Fig. 1 and Supplementary Table S2). Carnitines and a large portion of caffeine, urea, and bile acid pathways had strong positive PC_{LF} loadings, while histamine, dihydroxyoctadecamonoenoic acids (diHOMEs), glucocorticoid, and histamine metabolites were associated with worse LFTs (Fig. 1 and Supplementary Table S2). Additionally, based on the MetaboAnalyst metabolite set enrichment analyses, only the histidine metabolism pathway was identified as significantly associated, with a false discovery rate (FDR) = 0.0002 (Supplementary Table S3).

In addition to this primary analysis, we conducted supplementary analyses using a one-vs-rest strategy to examine each LFT group individually, spanning NG vs. rest (Supplementary Table S4), EG vs. rest (Supplementary Table S5), NG vs. rest (Supplementary Table S6), and NG + EG vs. rest (Supplementary Table S7). However, no significant associations (Q-value < 0.05) were observed with metabolite PCs.

Phenotypic and Clinical Characteristics	CAMP (n = 660)	GACRS (n = 1151)	P-value
Demographic variables			
Age [in years], Mean (SD)	13.51 (1.76)	9.22 (1.88)	<2e-16
Sex, n (%)			0.70
Female	262 (39.7%)	469 (40.7%)	
Male	398 (60.3%)	682 (59.3%)	
BMI [in kg/m^2], Mean (SD)	21.92 (4.81)	18.28 (3.77)	<2e-16
Race, n (%)			<2e-16
White (%)	445 (67.4%)	–	
Black (%)	93 (14.1%)	–	
Hispanic (%)	68 (10.3%)	1151 (100%)	
Other (%)	54 (8.2%)	–	
Treatment group, n (%)			–
Budesonide	195 (29.5%)	–	
Nedocromil	465 (70.5%)	–	
Placebo	0 (0.0%)	–	
Prior ICS Usage, n (%)			2e-8
Yes	245 (37.3%)	588 (51.1%)	
No	412 (62.7%)	563 (48.9%)	
Height [in cm], Mean (SD)	159.45 (11.77)	132.66 (11.85)	<2e-16
Clinical variables			
AHR, Mean (SD)			
PC_{20}	0.93 (1.57)	–	–
PD_{20}	–	1.93 (2.53)	–
FEV_1% Predicted, Mean (SD)	94.40 (14.20)	98.89 (17.15)	3e-8
FVC % Predicted, Mean (SD)	105.80 (12.85)	104.60 (16.51)	0.09
FEV_1/FVC, Mean (SD)	77.92 (9.04)	84.23 (7.90)	<2e-16
FEF_{25-75} [in L], Mean (SD)	2.47 (0.94)s	2.04 (0.74)	<2e-16
FEF_{25-75}/FVC, Mean (SD)	0.72 (0.23)	0.98 (0.31)	<2e-16
log_{10} (IgE) [in kU/L], Mean (SD)	2.65 (0.63)	2.50 (0.67)	4.4e-6
log_{10} (Eosinophil Count) [in kU/L], Mean (SD)	2.40 (0.49)	2.60 (0.41)	<2e-16
Neutrophil %, Mean (SD)	49.86 (10.90)	–	–

BMI = body mass index; CAMP = Childhood Asthma Management Program; GACRS = Genetics of Asthma in Costa Rica Study; SD = standard deviation; ICS = inhaled corticosteroids; AHR = airway hyperresponsiveness; n = number. P-value based on the t-test for continuous variables and the chi-square test for categorical variables.

Table 1: **Characteristics of CAMP and GACRS at the time of metabolomic profiling.**

Replication of metabolite PC_{LF} with asthma phenotypes in GACRS

PC_{LF} was recapitulated in GACRS using the CAMP PC_{LF} loadings. No significant associations were observed with any demographic variables, including age, sex and BMI (Table 3). Significant associations were observed between PC_{LF} and multiple lung function measures in GACRS: FEV_1/FVC (β = 0.59; P-value = 2.37×10^{-4}), FEF_{25-75} (β = ,0.03; P-value = 0.02), and FEF_{25-75}/FVC (β = 0.02; P-value = 1.75×10^{-4}) (via

Metabolite PC	NG mean (SD)	ED mean (SD)	RG mean (SD)	RG/ED mean (SD)	Coefficient (Std. Error)	P-value	Q-value
PC29 (PC_{LF})	0.40 (1.85)	−0.01 (1.82)	−0.14 (1.83)	−0.23 (1.93)	−0.13 (0.04)	6.04E-04	0.03
PC12	−0.50 (3.16)	−0.13 (2.81)	−0.18 (2.91)	0.43 (2.74)	0.06 (0.03)	0.01	0.41
PC31	−0.17 (1.86)	−0.22 (1.69)	0.29 (1.76)	0.11 (1.80)	0.09 (0.04)	0.02	0.45
PC57	0.15 (1.46)	0.03 (1.29)	−0.03 (1.31)	−0.14 (1.23)	−0.11 (0.05)	0.05	0.54
PC14	0.14 (2.75)	0.18 (2.70)	−0.29 (2.58)	−0.05 (2.96)	−0.06 (0.03)	0.05	0.54

PC = principal component; LFT = lung function trajectory; SD = standard deviation; BMI = body mass index; NG = normal growth; ED = early decline; RG = reduced growth; RG/ED = reduced growth with early decline; LFT groups categorize lung growth patterns from childhood to early adulthood into distinct classifications based on the severity of lung function decline and diminished lung growth throughout development, ordered from least to most severe as NG, ED, RG, and RG/ED.

Table 2: **Associations of metabolite PCs with LFT via ordinal logistic regression, adjusted for age, sex, BMI, race, and height (P-value < 0.05).**

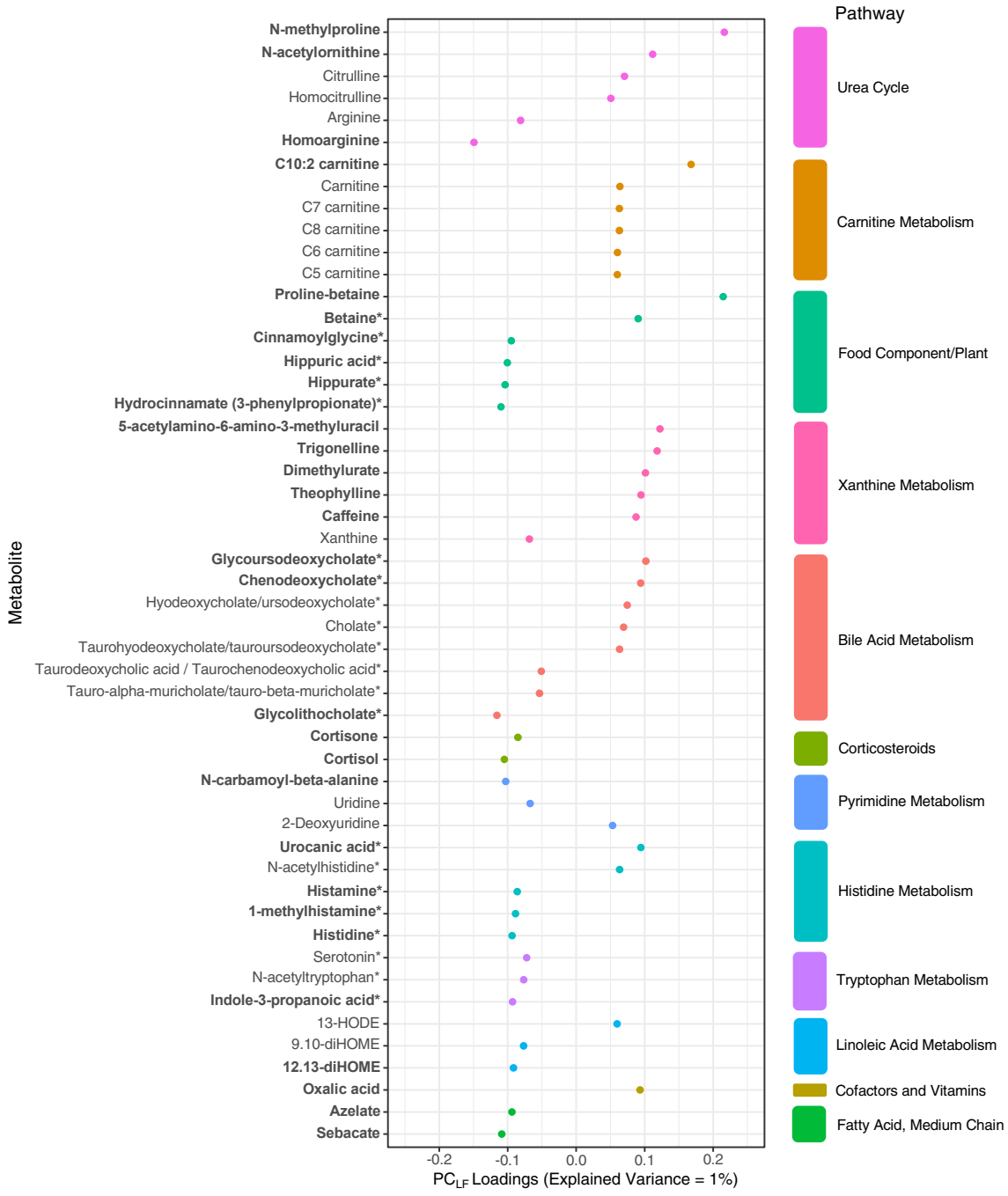

Fig. 1: **PC_{LF} principal component loading plots organized by pathway with an overall significance threshold of alpha = 0.20.** Bold metabolites signify a P-value < 0.05. * refers to microbial metabolites. Significance was determined using parallel analysis, a method for identifying significant loadings in a PC: alpha = 0.05 corresponds to an absolute loading greater than 0.0830, and alpha = 0.20 corresponds to an absolute loading greater than 0.0506. Positive loadings on PC_{LF} represent metabolites associated with improved lung function, while negative loadings indicate those linked to deteriorated lung function as PC_{LF} itself is associated with shifts towards better LFTs.

Phenotypic and Clinical Characteristics	CAMP		GACRS	
	Coefficient (Std. Error)	P-value	Coefficient (Std. Error)	P-value
Demographic variables				
Age [in years]	−0.01 (0.04)	0.89	0.02 (0.04)	0.63
Sex	0.001 (0.04)	0.98	0.04 (0.04)	0.36
BMI [in kg/m^2]	−0.10 (0.10)	0.35	−0.09 (0.08)	0.27
Race	−0.04 (0.05)	0.37	–	–
Treatment Group	−0.01 (0.05)	0.80	–	–
ICS Usage	−0.07 (0.04)	0.09	0.03 (0.04)	0.49
Height [in cm]	−0.41 (0.25)	0.10	0.002 (0.002)	0.34
Clinical variables				
AHR (PC_{20} in CAMP, PD_{20} in GACRS)	−0.01 (0.03)	0.68	−0.14 (0.06)	0.02
FEV_1% Predicted	0.77 (0.29)	0.008	0.15 (0.33)	0.64
FVC % Predicted	0.84 (0.26)	0.001	−0.61 (0.31)	0.05
FEV_1/FVC	0.03 (0.19)	0.89	0.59 (0.16)	2.37E-04
FEF_{25-75} [in L]	0.01 (0.02)	0.45	0.03 (0.01)	0.02
FEF_{25-75}/FVC	1.58 (4.93)	0.75	0.02 (0.01)	1.75E-04
log_{10} (IgE) [in kU/L]	0.01 (0.01)	0.52	−0.002 (0.01)	0.87
log_{10} (Eosinophil Count) [in kU/L]	0.01 (0.01)	0.23	−0.01 (0.01)	0.08
Neutrophil %	−0.61 (0.22)	0.005	−0.007 (0.003)	0.01

BMI = body mass index; CAMP = Childhood Asthma Management Program; GACRS = Genetics of Asthma in Costa Rica Study; SD = standard deviation; ICS = inhaled corticosteroids; AHR = airway hyperresponsiveness.

Table 3: **Association between PC_{LF} and demographic & clinical variables in CAMP and GACRS via generalized linear models, with clinical models adjusted for age, sex, BMI, race (race adjustment for CAMP only), and height (height adjustment for lung function only).**

GLMs adjusted for age, sex, BMI, race, and height). Additionally, higher levels of PC_{LF} were associated with decreased circulating neutrophil percentage in both CAMP (β = −0.61; P-value = 0.005) and GACRS (β = −0.007; P-value = 0.01) (via GLMs adjusted for age, sex, BMI, and race) (Table 3).

Association between metabolite PC_{LF}, miRs, and significant miR genetic targets

A microRNAome-wide association analysis of PC_{LF} and 252 baseline CAMP miRs passing quality control identified one statistically significant miR after multiple testing correction, miR-143-3p, where higher expression levels were associated with lower PC_{LF} (β = −0.08, SE = 0.02; P-value = 1.3 75 × 10^{-4}, Q = 0.03) (via GLMs adjusted for age, sex, BMI, and race), indicating an association with worse lung function outcomes (Table 4). We then identified 1064 gene targets (37,427 SNPs) of miR-143-3p based on the validated microRNA-target databases (miRecords, miRTarBase and TarBase).[29] Using an additive genetic model and meta-analyzing the results across GACRS and CAMP, three SNPs were significantly associated with PC_{LF} based on the ENT80 threshold: chr7_5901578_T_C found in *CCZ1* (β = 0.14, SE = 0.04; P-value = 2.6 × 10^{-5}); chr2_40180727_G_A found in *SLC8A1* (β = −0.16, SE = 0.04; P-value = 3.9 × 10^{-5}); and chr11_79047324_T_C found in *TENM4* (β = −0.22, SE = 0.05; P-value = 4.9 × 10^{-5}) (via GLMs adjusted for age, sex, and the first 4 genome-wide PCs) (Table 5).

Discussion

This study examined the relationship between metabolites and LFTs among children with asthma, with further genomic characterization using miRs and their gene targets. Using metabolite principal components and four ordinal LFTs,[4] we identified a significant principal component associated with improved LFT in CAMP; when we recapitulated this in GACRS we observed a significant increase in many spirometric measures with increasing PC_{LF}, substantiating our

miR	CAMP		GACRS		Meta-Analysis		
	Coefficient (Std. Error)	P-value	Coefficient (Std. Error)	P-value	Coefficient (Std. Error)	P-value	Q-value
miR 143.3p	−0.08 (0.03)	0.02	−0.09 (0.03)	0.002	−0.08 (0.02)	0.0001	0.03

miRs = microRNAs; BMI = body mass index; CAMP = Childhood Asthma Management Program; GACRS = Genetics of Asthma in Costa Rica Study.

Table 4: **Associations between miRs and PC_{LF} via generalized linear models, with adjustments for age, sex, BMI, and race (race adjustment specific to CAMP) (Q-value < 0.05).**

Gene	SNP	CAMP		GACRS		Meta-Analysis	
		Coefficient (Std. Error)	P-value	Coefficient (Std. Error)	P-value	Coefficient (Std. Error)	P-value
CCZ1	chr7_5901578_T_C	0.18 (0.06)	0.001	0.12 (0.04)	0.001	0.14 (0.03)	2.55E-05
SLC8A1	chr2_40180727_G_A	−0.22 (0.07)	0.001	−0.13 (0.05)	0.001	−0.16 (0.04)	3.86E-05
TENM4	chr11_79047324_T_C	−0.26 (0.08)	0.0009	−0.18 (0.07)	0.0009	−0.22 (0.05)	4.95E-05

P-value significance threshold for multiple corrections was determined based on ENT80, calculated as 0.05/792. miR = microRNA; SNPs = single nucleotide polymorphisms; eQTL = expression Quantitative Trait Loci; BMI = body mass index; PCs = principal components; CAMP = Childhood Asthma Management Program; GACRS = Genetics of Asthma in Costa Rica Study.

Table 5: **Associations between miR-143-3p gene targets SNPs and PC_{LF} via an eQTL analysis, with adjustments for age, sex, BMI, and the first 4 genome-wide PCs (P-value < 6E-05).**

initial finding. When considering potential genomic influences, we identified a significant association between PC_{LF} and baseline miR-143-3p, a miR that inhibits airway remodeling and has been identified as a therapeutic target for asthma.[32,33] SNPs in three miR-143-3p target genes (*CCZ1*, *SLC8A1*, and *TENM4*) were also significantly associated with PC_{LF}. Metabolites with strong loadings on the PC_{LF} include urea cycle, caffeine metabolites, corticosteroids, carnitines, and microbial metabolites from secondary bile acid, tryptophan, linoleate, and histidine metabolism pathways. Each of these metabolites has demonstrated significant biological implications for both lung inflammation and functional development, warranting further investigation.

Multiple metabolites in the urea cycle had large loadings on PC_{LF}, with N-methylproline having the largest loading of all metabolites and ornithine, citrulline, and arginine metabolites also contributing substantially. These metabolites have consistently been identified as associated with asthma phenotypes, and as discriminators of asthma diagnosis in large metabolomics studies.[34] It has been demonstrated that arginine metabolism is a key regulator of nitric oxide (NO) and, subsequently, the development of lung inflammation, including pro-inflammatory processes.[35,36] However, randomized trials of arginine supplementation in severe asthma did not demonstrate a significant reduction in exacerbations. In fact, this study found that higher citrulline levels and a lower arginine availability index were more efficacious, with higher exhaled NO levels and lower exacerbation events.[37] This study observed consistent findings to this for citrulline and arginine, suggesting that this same relationship may also be important for long-term lung function.

Multiple exogenous metabolites evidenced high loadings on PC_{LF}, which is of particular interest because these exposures may point towards modifiable approaches to optimizing LFTs. Notably, caffeine metabolites were associated with improved LFTs. There is a strong biological rationale for this pathway, because a primary derivative of caffeine is theophylline; an approved asthma medication known for its positive effects on bronchodilation and suppression of airway inflammation.[38] Increased caffeine intake has been suggested as a form of self-treatment for asthma.[39,40] Though to confirm an increased caffeine intake, we would require the nutritional information of the patients, which, unfortunately, is not available in either the CAMP or GACRS. Regardless, these findings do suggest that the potential benefits of caffeine intake may extend beyond acute asthma treatment to potential long-term benefits on lung function. Other well-known coffee metabolites, including trigonelline[41] and gut microbial-derived metabolites of polyphenols (hydrocinnamate, hippurate, and cinnamoylglycine),[42,43] also contributed positively to PC_{LF}.

The association between reduced endogenous corticosteroids and worse LFTs is not surprising from a physiological perspective. Reduced corticosteroids have been identified as a hallmark sign of asthma with further adrenal suppression of endogenous corticosteroids resulting from inhaled corticosteroid treatment.[44] In this study, we identified reduced cortisol and cortisone associated with worse LFT. While ICS is often confounded in studying this association between any metric of disease severity and disease pathophysiology, in using CAMP for this study, we benefit from the RCT design with ICS that circumvents this potential confounding by asthma severity. Therefore, the observation of decreased corticosteroids with worse LFT remains substantiated as a feature of asthma physiology. This study was limited in the assessment of steroid classes, as cortisone and cortisol were the only measured steroids.

Other strong contributors to PC_{LF} included carnitines, diHOMES, and potential microbial-derived metabolites from bile acid, histidine, and tryptophan metabolism. Recent research has observed decreased carnitines in association with severe asthma.[45] We also identified that increased carnitines were associated with improved LFTs. Another compelling finding was the negative loading of 9,10-diHOME and 12,13-diHOME on PC_{LF}. Both diHOMEs can be produced by gut bacteria and affect immune intolerance and are known asthma risk factors.[46] Severe asthma metabo-endotypes have also demonstrated increased levels of diHOMES compared with other metabo-endotypes.[19] Previous

research has demonstrated that conjugated bile acids have a protective effect on asthma via inhibiting unfolded protein response (UPR) transducers that attenuate allergen-induced airway inflammation.[47] Our findings suggest that there are both positive and negative weightings of bile acids on PC_{LF}, which may be attributed to the complex interplay between primary and secondary bile acid metabolism pathways. Further research is needed to clarify if the mechanistic effect on conjugated bile acids may influence long-term lung function. Metabolites involved in histidine metabolism were also represented in PC_{LF}, with histamine metabolites being strong drivers of worse LFTs, suggesting that those with worse LFTs also have higher histamine levels, which should not be surprising as these individuals likely have a more severe allergic phenotype. Furthermore, as the only pathway that was found to be significantly associated with PC_{LF} (via the MetaboAnalyst metabolite set enrichment analyses), it underscores the importance of histidine metabolism. Histamine, a prominent product of histidine metabolism, has long been recognized for its central role in allergic reactions and asthma pathophysiology.[48,49] Recent studies indicate a rise in microbes that secrete histamine in the gut of individuals with asthma.[50] Such microbes appear to have a discernible impact on lung immune responses.[51] Considering that shifts in microbial composition and metabolism during infancy can shape the risk of asthma in childhood,[52,53] it underscores the importance of delving deeper into histidine metabolism and its potential microbial roots in severe childhood asthma cases. Finally, we observed an inverse relationship between tryptophan metabolites and PC_{LF}, which further validates the literature suggesting increases in tryptophan metabolites are associated with increased severity.[54,55]

While several of these metabolites have previously been associated with other asthma phenotypes, this study identified a composite measure of metabolites that is associated with long term lung function. While many of the metabolites contributing to LFT have concrete biological links to asthma, several of them may also provide clues into potential modifications that may alter adverse LFTs. While further work is certainly merited to substantiate these metabolite pathways, it is still worth considering their potential to provide simple, long-term solutions for improving overall health outcomes. This is a critical point to consider, given that the overall burden from asthma increases as lung function continues to decline with age, including an increased risk of chronic obstructive pulmonary disease.[56] Any intervention resulting in incremental improvement in LFT may substantially improve overall health outcomes. Particularly, the significant role of exogenous metabolites, such as caffeine derivatives, in optimizing LFT points to feasible, non-invasive interventions that could be beneficial in real-world settings. Likewise, the notable associations we observed between specific endogenous metabolites and LFTs underline the intricate interplay of our body's internal biochemical processes. The balance of these endogenous metabolites is crucial as they are not only markers of metabolic activity but also regulators of various physiological processes. For instance, disruptions in the urea cycle, as indicated by the levels of ornithine, citrulline, and arginine, can influence nitric oxide production, which in turn can modulate inflammation and bronchial responsiveness, both of which are critical in asthma pathophysiology.[57] Similarly, alterations in tryptophan metabolism can affect serotonin levels, which have been linked to bronchoconstriction and immune responses.[58–60] Thus, understanding and maintaining optimal levels of these endogenous metabolites can potentially offer avenues for therapeutic interventions, aiming to restore the body's natural equilibrium and, in turn, improve lung function and overall asthma outcomes.

MiRs are small noncoding RNAs that play an important role in regulating gene expression for specific target genes. There was one significant miR associated with PC_{LF}, miR-143-3p. Mechanistic studies have demonstrated that miR-143-3p is functionally relevant by inhibiting airway remodeling and potentially modulating inflammatory or pro-inflammatory processes.[32,33] In the present study, however, higher levels of miR-143-3p were associated with worse LFTs, a finding that appears to challenge its mechanistically understood role. One plausible explanation could be a compensatory increase in miR-143-3p levels as a protective response to deteriorating lung function. Alternatively, while miR-143-3p has been known to inhibit airway remodeling, it's possible that in certain microenvironments or stages of asthma progression, its role might diverge from its typical anti-inflammatory function. Furthermore, interactions with other molecules or cellular pathways might influence how miR-143-3p affects LFTs. This underscores the importance of studying miRNAs within specific physiological contexts and highlights the intricate balance and fine-tuning involved in gene regulation and its impact on disease outcomes. Future investigations are warranted to further elucidate the nuances of miR-143-3p's role in asthma and its connection with lung function.

Furthermore, SNPs in miR-143-3p gene targets (*CCZ1, SLC8A1, TENM4*) associated with PC_{LF} also have a strong biologic rationale for their involvement in asthma; *SLC8A1* has been implicated in asthma and exhaled NO levels previously,[61] and plasma concentration of *TENM4*-a protein involved in neural development-was found to be a risk factor for hay fever, allergic rhinitis, and asthma,[62] there is no clear relationship between *CCZ1* and asthma. The association of certain SNPs in the gene targets of miR-143-3p with PC_{LF} adds another layer of complexity to our understanding of the genetic underpinnings of asthma and

lung function. Specifically, the involvement of *SLC8A1* in asthma and its association with exhaled NO levels suggest that it may play a pivotal role in airway inflammation, a hallmark of asthma. Additionally, the link between plasma concentrations of *TENM4*—a protein integral to neural development—and allergic conditions hints at the intricate web of physiological processes that might converge to influence asthma phenotypes. While the exact mechanisms remain to be fully elucidated, these findings underscore the interconnectedness of seemingly distinct biological pathways in determining respiratory health outcomes. Interestingly, while *CCZ1* is a gene target of miR-143-3p, its direct relationship with asthma remains ambiguous. This highlights the diversity in function and impact of miR target genes and prompts further exploration to discern the potential role, if any, of *CCZ1* in asthma pathogenesis.

There are several strengths and limitations of this study. This study looked at the association between metabolites and LFTs among children with asthma, as opposed to single time-point LF measures, with further characterization using miR and genetics. Second, this analysis includes a diverse range of children, including White, Black, and Hispanic. Therefore, the findings from this study may generalize to several populations. To date, no studies have examined the relationship between both metabolomics and miRs together among children with asthma, another point of innovation. However, there are several limitations that should be considered. While we employed the GACRS cohort to replicate the metabolite PC_{LF} findings and observed consistent associations with alternative asthma phenotypes, the necessity of a cohort with LFT data is paramount. Currently, there's an absence of such a cohort, underscoring the need to establish a large cohort to validate the findings of this study. The metabolomic profiling utilized relative quantification for most of the metabolites that were profiled, as is the case in the majority of global metabolomics profiling platforms. The miRs profiling was performed at a different time point in CAMP; however prior work has shown miRs have long-term effects on asthma and lung function.[20,63] Future research that employs targeted metabolomic analyses on identified LFT-related metabolites and their pathways, and simultaneously profile miRs, is necessary to better elucidate their relationship with miR-143-3p. Additionally, while our study identified miR-143-3p as a microRNA associated with PC_{LF}, the exact functional role of this miRNA remains to be confirmed through advanced experimentation. Subsequent studies employing targeted functional assays will be crucial in determining the precise function of miR-143-3p in relation to LFT. Finally, to address the high degree of collinearity commonly seen in metabolomics data, our analytic approach for this primary analysis into LFTs and metabolomics incorporated data reduction using PCA, a standard approach for metabolomics data.[64,65] Given that machine learning methods like neural networks and random forests might offer more nuanced and comprehensive interpretations,[66,67] research towards utilizing alternative non-linear data reduction approach for future studies is warranted.

This study identified a PC_{LF} using metabolomics profiling that provides valuable information on physiology and insight into potential modifications that may improve overall LFTs. We further identify miR-143-3p as associated with PC_{LF}, further implicating pathways of inflammation.

Contributors

JLS and KM contributed to conceptualization of the study; MH performed the quality control; KM performed the primary statistical data analyses. NP, PK, RSK, and KM contributed to the downstream analyses. CC and RG generated the metabolomic data for TOPMed. JLS, SB, and MM verified the underlying data. DB, JLS, CEW, KM, NP, RSK, SNR, and MM contributed to the original draft preparation. All authors reviewed, provided feedback on, and approved the final manuscript.

Data sharing statement

The genetic, metabolomic, and phenotype data supporting the findings of this study can be accessed openly in the database of Genotypes and Phenotypes (dbGaP) under the following accessions: CAMP Study (Accession: phs001726.v2.p1) and GACRS Study (Accession: phs000988.v5.p1). Additionally, microRNA data can be accessed openly in the Gene Expression Omnibus (GEO) with the respective accessions: CAMP Study (Accession: GSE134897) and GACRS Study (Accession: GSE244573).

Declaration of interests

JLS is a scientific advisor to Precion, Inc., receives grants and consulting fees from TruDiagnostic and Ahara Corp, and holds patents with TruDiagnostic. KLS is employed by Vertex Pharmaceuticals. STW is a board member of Histolix and receives royalties from UpToDate. All other authors declare no potential, perceived, or real conflict of interest regarding the content of this manuscript.

Acknowledgements

Molecular data for the Trans-Omics in Precision Medicine (TOPMed) program was supported by the National Heart, Lung, and Blood Institute (NHLBI). Metabolomics for "NHLBI TOPMed: Childhood Asthma Management Program (CAMP)" (phs001726) and "NHLBI TOPMed: The Genetic Epidemiology of Asthma in Costa Rica (CRA)" (phs000988) was performed at Broad Institute and Beth Israel Metabolomics Platform (HHSN268201600034I). Core support including phenotype harmonization, data management, sample-identity QC, and general program coordination were provided by the TOPMed Data Coordinating Center (R01HL-120393; U01HL-120393; contract HHSN268201800001I). Current work was supported by NIH NHLBI grants R01 HL139634, R01 HL155742 and K99HL159234.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2024.105025.

References

1 Network TGA. *The global asthma report*. Auckland: New Zealand; 2018.
2 Collaborators GBDCRD. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir Med*. 2017;5(9):691–706.

3 Nurmagambetov T, Kuwahara R, Garbe P. The economic burden of asthma in the United States, 2008-2013. *Ann Am Thorac Soc*. 2018;15(3):348–356.
4 McGeachie MJ, Yates KP, Zhou X, et al. Patterns of growth and decline in lung function in persistent childhood asthma. *N Engl J Med*. 2016;374(19):1842–1852.
5 Dunn WB, Broadhurst DI, Atherton HJ, Goodacre R, Griffin JL. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chem Soc Rev*. 2011;40(1):387–426.
6 Kelly RS, Dahlin A, McGeachie MJ, et al. Asthma metabolomics and the potential for integrative omics in research and the clinic. *Chest*. 2017;151(2):262–277.
7 Pite H, Morais-Almeida M, Rocha SM. Metabolomics in asthma: where do we stand? *Curr Opin Pulm Med*. 2018;24(1):94–103.
8 Schoettler N, Strek ME. Recent advances in severe asthma: from phenotypes to personalized medicine. *Chest*. 2020;157(3):516–528.
9 Desai M, Oppenheimer J. Elucidating asthma phenotypes and endotypes: progress towards personalized medicine. *Ann Allergy Asthma Immunol*. 2016;116(5):394–401.
10 Decrue F, Gorlanova O, Usemann J, Frey U. Lung functional development and asthma trajectories. *Semin Immunopathol*. 2020;42(1):17–27.
11 Lee A, Kim SN. Serum MicroRNA on inflammation: a literature review of mouse model studies. *Biomarkers*. 2020;25(7):513–524.
12 Hough KP, Curtiss ML, Blain TJ, et al. Airway remodeling in asthma. *Front Med*. 2020;7:191.
13 Sharma R, Tiwari A, McGeachie MJ. Recent miRNA research in asthma. *Curr Allergy Asthma Rep*. 2022;22(12):231–258.
14 Thomsen SF. Genetics of asthma: an introduction for the clinician. *Eur Clin Respir J*. 2015;2. https://doi.org/10.3402/ecrj.v2.24643.
15 Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. *Immunol Rev*. 2011;242(1):10–30.
16 Toskala E, Kennedy DW, eds. *Asthma risk factors. International forum of allergy & rhinology*. Wiley Online Library; 2015.
17 The childhood asthma management program (CAMP): design, rationale, and methods. Childhood asthma management program research group. *Control Clin Trials*. 1999;20(1):91–120.
18 Kelly RS, Virkud Y, Giorgio R, Celedon JC, Weiss ST, Lasky-Su J. Metabolomic profiling of lung function in Costa-Rican children with asthma. *Biochim Biophys Acta Mol Basis Dis*. 2017;1863(6):1590–1595.
19 Kelly RS, Mendez KM, Huang M, et al. Metabo-endotypes of asthma reveal differences in lung function: discovery and validation in two TOPMed cohorts. *Am J Respir Crit Care Med*. 2021;205(3):288–299.
20 Tiwari A, Li J, Kho AT, et al. COPD-associated miR-145-5p is downregulated in early-decline FEV(1) trajectories in childhood asthma. *J Allergy Clin Immunol*. 2021;147(6):2181–2190.
21 Rozowsky J, Kitchen RR, Park JJ, et al. exceRpt: a comprehensive analytic platform for extracellular RNA profiling. *Cell Syst*. 2019;8(4):352–357.e3.
22 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
23 Joo J, Mak ACY, Xiao S, et al. Genome-wide association study in minority children with asthma implicates DNAH5 in bronchodilator responsiveness. *Sci Rep*. 2022;12(1):12514.
24 Reinke SN, Chaleckis R, Wheelock CE. Metabolomics in pulmonary medicine: extracting the most from your data. *Eur Respir J*. 2022;60(2):2200102.
25 Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*. 2001;125(1–2):279–284.
26 Franklin SB, Gibson DJ, Robertson PA, Pohlmann JT, Fralish JS. Parallel analysis: a method for determining significant principal components. *J Veg Sci*. 1995;6(1):99–106.
27 Peres-Neto PR, Jackson DA, Somers KM. Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology*. 2003;84(9):2347–2363.
28 Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 2021;49(W1):W388–W396.
29 Ru Y, Kechris KJ, Tabakoff B, et al. The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res*. 2014;42(17):e133.
30 Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet*. 2004;74(4):765–769.
31 Li M-X, Yeung JMY, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 2012;131(5):747–756.
32 Cheng W, Yan K, Xie LY, et al. MiR-143-3p controls TGF-β1-induced cell proliferation and extracellular matrix production in airway smooth muscle via negative regulation of the nuclear factor of activated T cells 1. *Mol Immunol*. 2016;78:133–139.
33 Cai XJ, Huang LH, Zhu YK, Huang YJ. LncRNA OIP5-AS1 aggravates house dust mite-induced inflammatory responses in human bronchial epithelial cells via the miR-143-3p/HMGB1 axis. *Mol Med Rep*. 2020;22(6):4509–4518.
34 Kelly RS, McGeachie MJ, Lee-Sarwar KA, et al. Partial least squares discriminant analysis and bayesian networks for metabolomic prediction of childhood asthma. *Metabolites*. 2018;8(4):68.
35 Benson RC, Hardy KA, Morris CR. Arginase and arginine dysregulation in asthma. *J Allergy*. 2011;2011:736319.
36 Zimmermann N, Rothenberg ME. The arginine-arginase balance in asthma and lung inflammation. *Eur J Pharmacol*. 2006;533(1–3):253–262.
37 Liao SY, Showalter MR, Linderholm AL, et al. l-Arginine supplementation in severe asthma. *JCI Insight*. 2020;5(13):e137777.
38 Rabe KF, Dent G. Theophylline and airway inflammation. *Clin Exp Allergy*. 1998;28(Suppl 3):35–41.
39 Welsh EJ, Bara A, Barley E, Cates CJ. Caffeine for asthma. *Cochrane Database Syst Rev*. 2010;2010(1):Cd001112.
40 Han YY, Forno E, Celedón JC. Urinary caffeine and caffeine metabolites, asthma, and lung function in a nationwide study of U.S. adults. *J Asthma*. 2022;59(11):2127–2134.
41 Hang D, Zeleznik OA, He X, et al. Metabolomic signatures of long-term coffee consumption and risk of type 2 diabetes in women. *Diabetes Care*. 2020;43(10):2588–2596.
42 Wikoff WR, Anfora AT, Liu J, et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci U S A*. 2009;106(10):3698–3703.
43 Ogawa M, Suzuki Y, Endo Y, Kawamoto T, Kayama F. Influence of coffee intake on urinary hippuric acid concentration. *Ind Health*. 2011;49(2):195–202.
44 Kachroo P, Stewart I, Kelly R, et al. The systematic use of metabolomic epidemiology, biobanks, and electronic medical records for precision medicine initiatives in asthma: findings suggest new guidelines to optimize treatment. *Nat Portfolio*. 2022. https://doi.org/10.21203/RS.3.RS-268507/V1.
45 Reinke SN, Naz S, Chaleckis R, et al. Urinary metabotype of severe asthma evidences decreased carnitine metabolism independent of oral corticosteroid treatment in the U-BIOPRED study. *Eur Respir J*. 2022;59(6):2101733.
46 Levan SR, Stamnes KA, Lin DL, et al. Elevated faecal 12,13-diHOME concentration in neonates at high risk for asthma is produced by gut bacteria and impedes immune tolerance. *Nat Microbiol*. 2019;4(11):1851–1861.
47 Nakada EM, Bhakta NR, Korwin-Mihavics BR, et al. Conjugated bile acids attenuate allergen-induced airway inflammation and hyperresponsiveness by inhibiting UPR transducers. *JCI Insight*. 2019;4(9):e98101.
48 Yamauchi K, Ogasawara M. The role of histamine in the pathophysiology of asthma and the clinical efficacy of antihistamines in asthma therapy. *Int J Mol Sci*. 2019;20(7):1733.
49 White MV, Slater JE, Kaliner MA. Histamine and asthma. *Am Rev Respir Dis*. 1987;135(5):1165–1176.
50 Barcik W, Pugin B, Westermann P, et al. Histamine-secreting microbes are increased in the gut of adult asthma patients. *J Allergy Clin Immunol*. 2016;138(5):1491–1494.e7.
51 Barcik W, Pugin B, Brescó MS, et al. Bacterial secretion of histamine within the gut influences immune responses within the lung. *Allergy*. 2019;74(5):899–909.
52 Arrieta MC, Stiemsma LT, Dimitriu PA, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med*. 2015;7(307):307ra152.
53 Smits HH, Hiemstra PS, Da Costa CP, et al. Microbes and asthma: opportunities for intervention. *J Allergy Clin Immunol*. 2016;137(3):690–697.
54 van der Sluijs KF, van de Pol MA, Kulik W, et al. Systemic tryptophan and kynurenine catabolite levels relate to severity of rhinovirus-induced asthma exacerbation: a prospective study with a parallel-group design. *Thorax*. 2013;68(12):1122–1130.

55 Gostner JM, Becker K, Kofler H, Strasser B, Fuchs D. Tryptophan metabolism in allergic disorders. *Int Arch Allergy Immunol*. 2016;169(4):203–215.
56 Lange P, Celli B, Agustí A, et al. Lung-function trajectories leading to chronic obstructive pulmonary disease. *N Engl J Med*. 2015;373(2):111–122.
57 Prado CM, Martins MA, Tibério IF. Nitric oxide in asthma physiopathology. *ISRN Allergy*. 2011;2011:832560.
58 Barnes PJ. Histamine and serotonin. *Pulm Pharmacol Therapeut*. 2001;14(5):329–339.
59 Nau F Jr, Miller J, Saravia J, et al. Serotonin 5-HT_2 receptor activation prevents allergic asthma in a mouse model. *Am J Physiol Lung Cell Mol Physiol*. 2015;308(2):L191–L198.
60 Herr N, Bode C, Duerschmied D. The effects of serotonin in immune cells. *Front Cardiovasc Med*. 2017;4:48.
61 Liu K, Liu Z, Qi H, et al. Genetic variation in SLC8A1 gene involved in blood pressure responses to acute salt loading. *Am J Hypertens*. 2018;31(4):415–421.
62 Li A, Liao W, Xie J, Song L, Zhang X. Plasma proteins as occupational hazard risk monitors for populations working in harsh environments: a mendelian randomization study. *Front Public Health*. 2022;10:852572.
63 McGeachie MJ, Davis JS, Kho AT, et al. Asthma remission: predicting future airways responsiveness using an miRNA network. *J Allergy Clin Immunol*. 2017;140(2):598–600.e8.
64 Worley B, Powers R. Multivariate analysis in metabolomics. *Current metabolomics*. 2013;1(1):92–107.
65 Lamichhane S, Sen P, Dickens AM, Hyötyläinen T, Orešič M. An overview of metabolomics data analysis: current tools and future perspectives. *Compr Anal Chem*. 2018;82:387–413.
66 Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science*. 2006;313(5786):504–507.
67 Ziegler A, König IR. Mining data with random forests: current options for real-world applications. *Wiley Interdiscip Rev*. 2014;4(1):55–63.